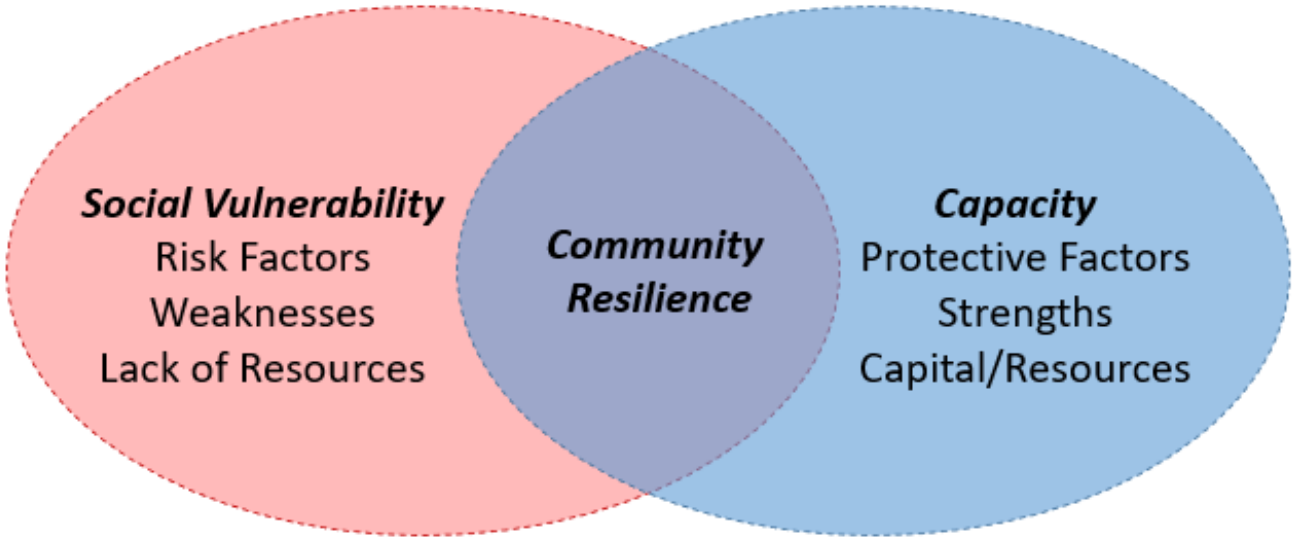
Small Area Modeling: A Tool for Providing Estimates for Disasters

Presented by: Bethany DeSalvo, PhD
Chief, Small Area Modeling & Development
Social, Economic, and Housing Statistics Division
U.S. Census Bureau

NOTE: This presentation is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the presenters and not necessarily those of the U.S. Census Bureau. The U.S. Census Bureau reviewed this data product for release. CBDRB-FY22-POP001-0132

Community Resilience & Social Vulnerability

- Social vulnerability is the risk of hazards to the physical and socially built environment, while community resilience is the capacity of individuals and households to absorb the stresses from a disaster.[2]
  - To eliminate the need to classify characteristics of an area as contributing to either vulnerability or resilience, resilience and vulnerability are viewed to represent two sides of the same resilience coin.[3]
- Social vulnerability mapping strengthens community resilience and reduces inequities.[1]
  - By helping communities better anticipate, respond, resist, and recover from disasters.



1 Van Zandt, Peacock, Henry, Grover, Highfield and Brody 2012; 2 Masterson, Jamie Hicks, Walter Gillis Peacock, Shannon Zandt, Himanshu Grover, Lori Field Schwarz, and John Cooper 2014; 3 Summers, Smith, Harwell, and Buck 2017

Community Resilience Estimates

Pandemic Uses

- Information from the CRE could be utilized by decision makers to:
  - Inform vaccine distribution.
  - Identify where to provide education on public health standards.
  - Pinpoint areas that are at greatest risk of inequitable outcomes from health risks such as pandemics.

Natural Disaster Uses

- Information from the CRE could be used by decision makers to:
  - Identify communities most incapable of rebuilding after a disaster.
  - Ensure response to disaster response is equitable.

Community Resilience Estimates (CRE)

- KEY BACKGROUND
  - Released in June 2020 in response to COVID-19 pandemic.
  - Uses **restricted microdata** unavailable outside of the Census Bureau.
  - Estimates are more **timely, accurate, reliable and granular** than anything else available.

    Uses 1-year ACS instead of 5-year ACS.
    Uses small area modeling to reduce sampling error by 30%-60%.
    Complete coverage at the tract level.

Metric

- Defined by the number of risk factors that make a community vulnerable (out of 10 possible risk factors).

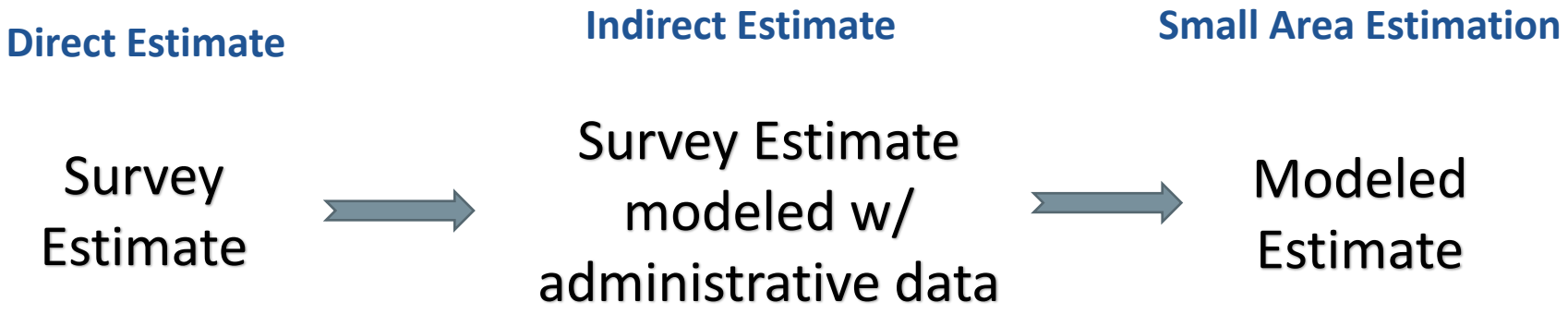
- CRE provides the number and percent of population with the following total number of risk factors for each geography.
  - ❑ Low: 0 risk factors
  - ❑ Medium: 1-2 risk factors
  - ❑ High: 3-10 risk factors

Small Area Estimation

- A small area is any geographic level or domain not supported by existing survey samples because of survey design.
  - Estimates from these small samples can yield high standard errors.
- SAE techniques "borrow strength" from other areas, such as administrative data sources, as well as other survey estimates.
- SAE techniques enable the Census Bureau to provide high quality estimates when direct methods cannot estimate quantities with acceptable accuracy.

Small Area Estimation

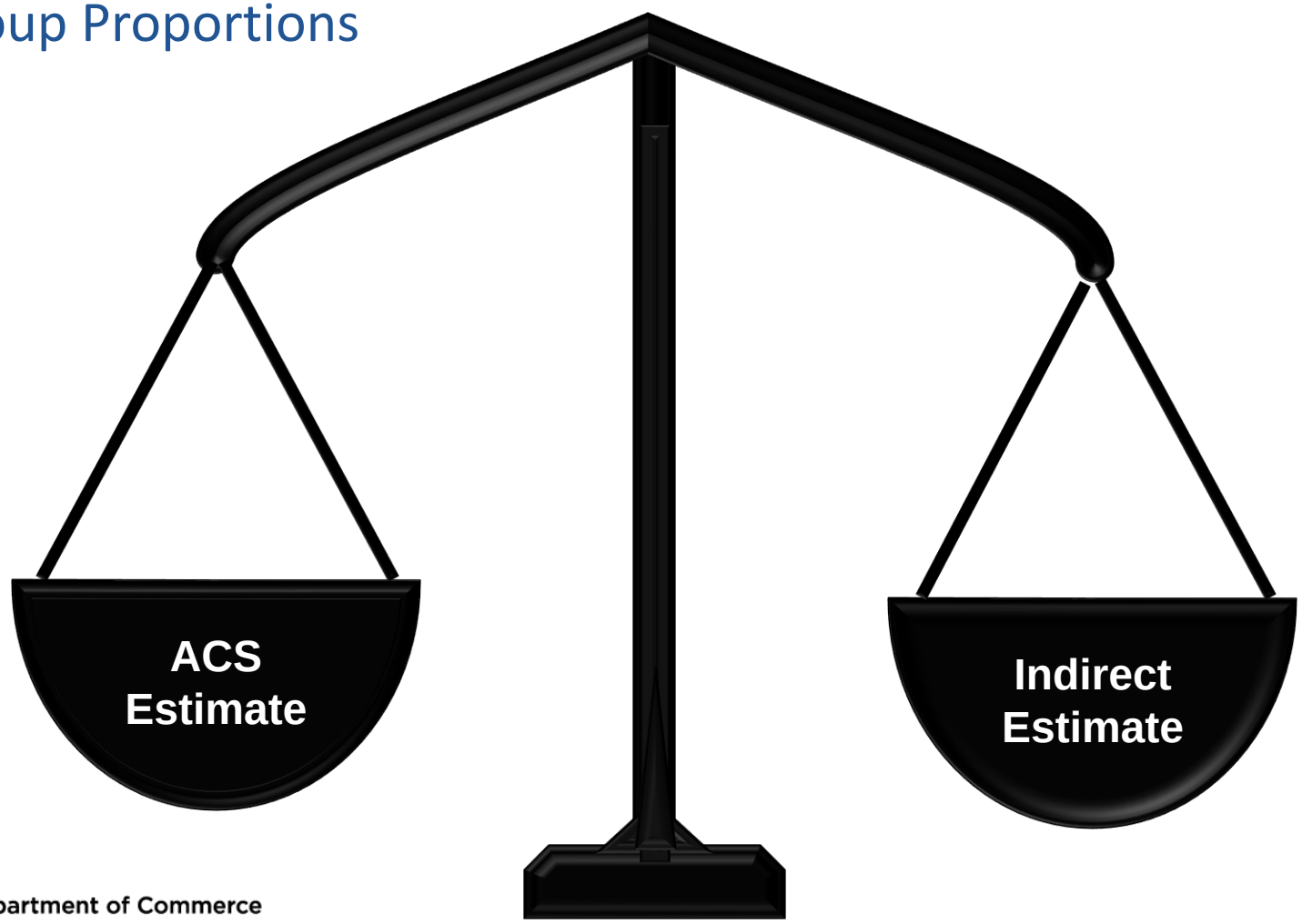
Small area statistical models take the **direct (survey)** estimate and **borrow strength** from related geographies, times, and/or groups through auxiliary data such as census data and administrative records.



Community Resilience Estimates

Weighted average between survey estimate and indirect estimate: Risk Factor Group Proportions

Proportion of population with 0 risk factors, 1-2 risk factors, and 3-10 risk factors tabulated from ACS at geography of interest



- Risk factor group proportions from poststratification
- Joint distribution of risk factor prevalence estimated at higher levels of geography (Census division by increasing levels of urbanicity) and by demographic groups (Age, race, and Hispanic origin) applied to tract population estimates

Community Resilience Estimates

ACS Direct Estimate: Risk Factor Group Proportions

Procedure

Tabulate risk factor groups (proportion, sampling variance) by tract in 1-year ACS

- 0 Risk Factors
- 1-2 Risk Factors
- 3 or more Risk Factors

Smooth sampling variance using a *generalized variance function* (gvf)

Risk Factors

1. Income-to-Poverty Ratio less than 130 percent
2. Single or Zero Caregiver Household (only one or no individuals who are 18-64)
3. Crowding (more than 0.75 persons per room)
4. Communication Barrier (limited English-speaking households or no one over the age of 16 with high school diploma)
5. Households without Full-time, Year-round Employment
6. Disability
7. No Health Insurance
8. Age 65+
9. No Vehicle Access
10. No Broadband Internet Access

Community Resilience Estimates

Indirect Estimate: Risk Factor Group Proportions

Procedure

Apply distribution of complete cross-classification of risk factors obtained at higher levels of geography and demographic groups to tract-level population estimates

Aggregate tract estimates with risk factor detail to broad risk factor group proportions

- 0 Risk Factors
- 1-2 Risk Factors
- 3 or more Risk Factors

Calculate model error variance by averaging over squared differences between county indirect and direct estimates

Cross-classification of Risk Factors:

0 = does not have risk factor
1= has risk factor

Risk Factor 1	**Risk Factor 2**	**….**	**Risk Factor 9**	**Risk Factor 10**	Probability
Income-to-poverty less than 130 percent	Single or 0 Caregiver	….	No Vehicle Access	No Broadband	
0	0	….	0	0	p1
0	0	….	0	1	p2
0	0	….	1	0	p3
.	.	….	.	.	.
.	.	….	.	.	.
1	1	1	1	1	p1024

Community Resilience Estimates

SAE modeled estimate: Weighted average between survey estimate and indirect estimate

Procedure

Calculate weights based on the sampling variance of ACS direct proportions by risk factor group and the model error variance from the indirect estimate

Derive SAE modeled proportions for each risk group from weighted proportions

Apply SAE modeled proportions to population estimates to get the modeled counts in each risk factor group

Weight Calculation

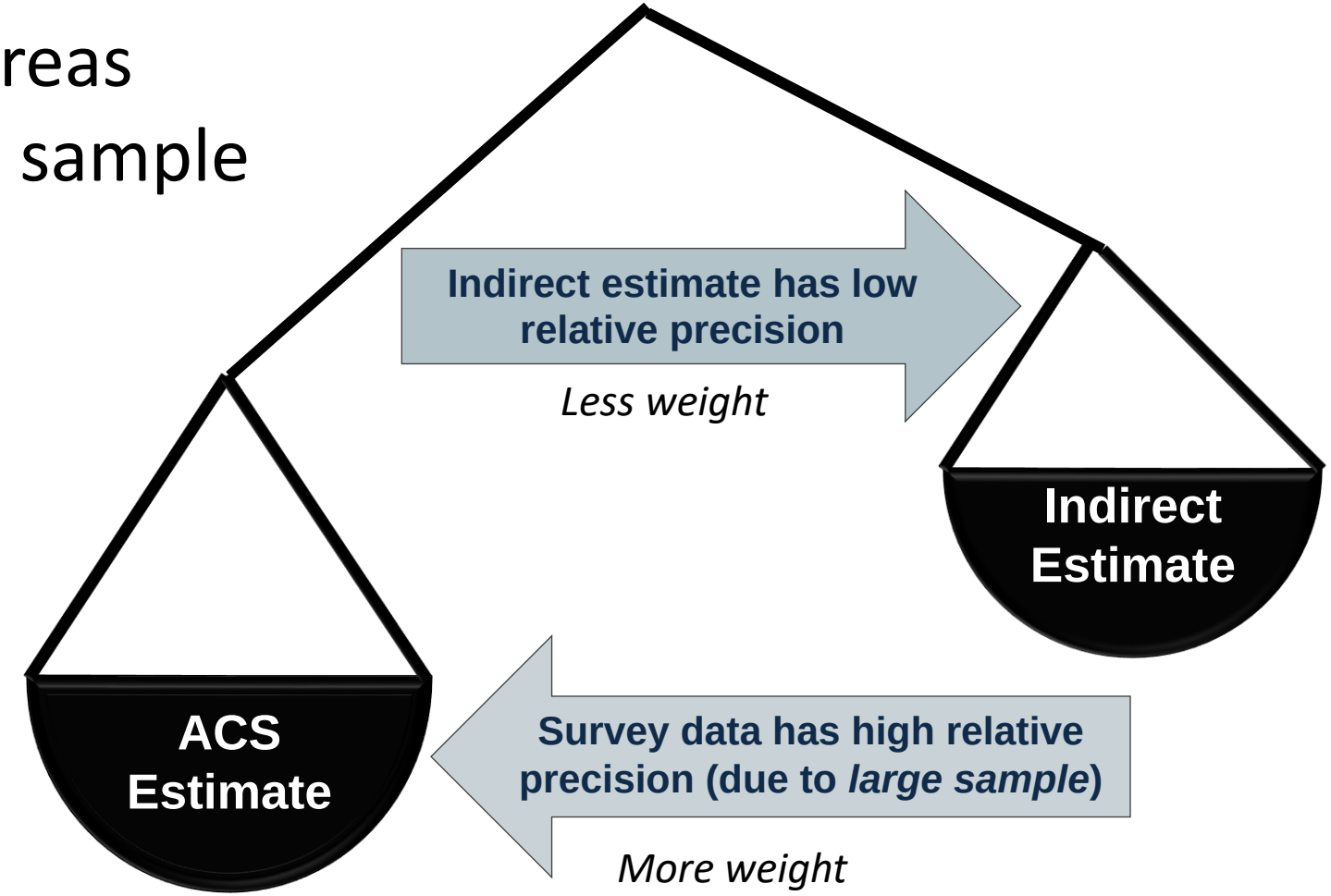
$$\text{Weight} = \frac{\text{indirect estimate variance}}{(\text{direct estimate variance} + \text{indirect estimate variance})}$$

SAE Modeled Proportions

(Weight) x (direct proportion) +
(1-Weight) x (indirect proportion)

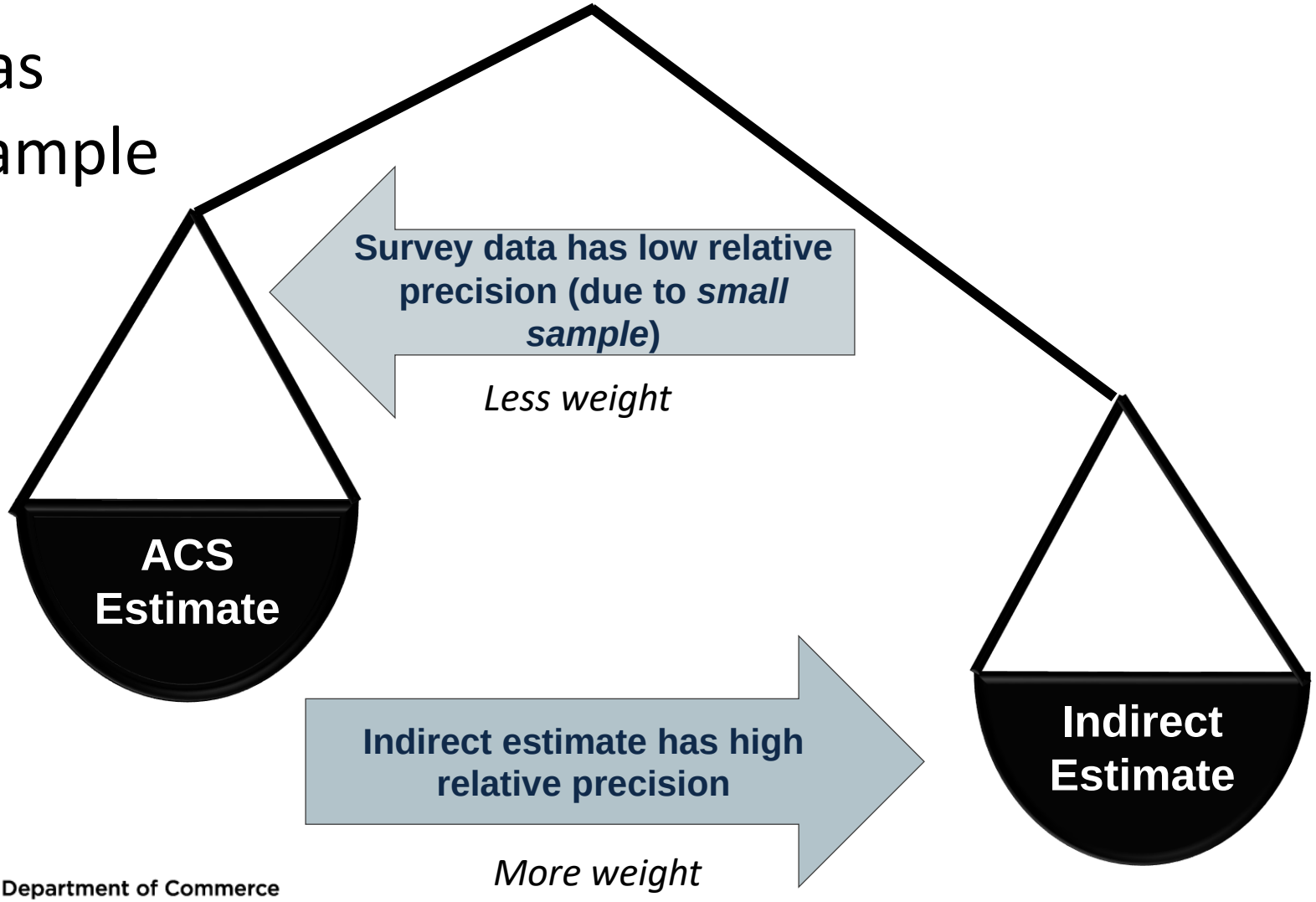
Community Resilience Estimates

SAEs for areas with large sample sizes



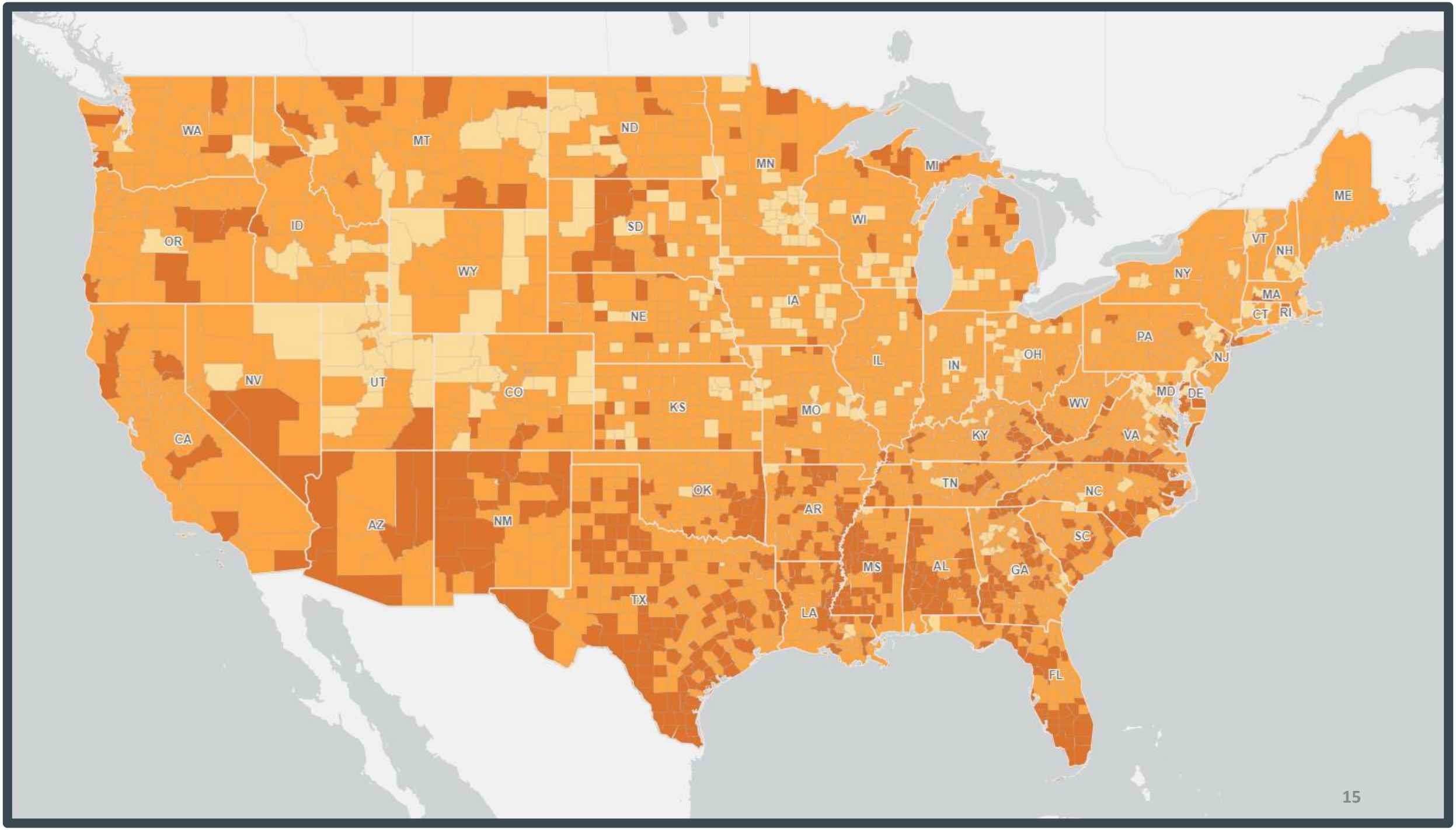
Community Resilience Estimates

SAEs for areas with small sample sizes

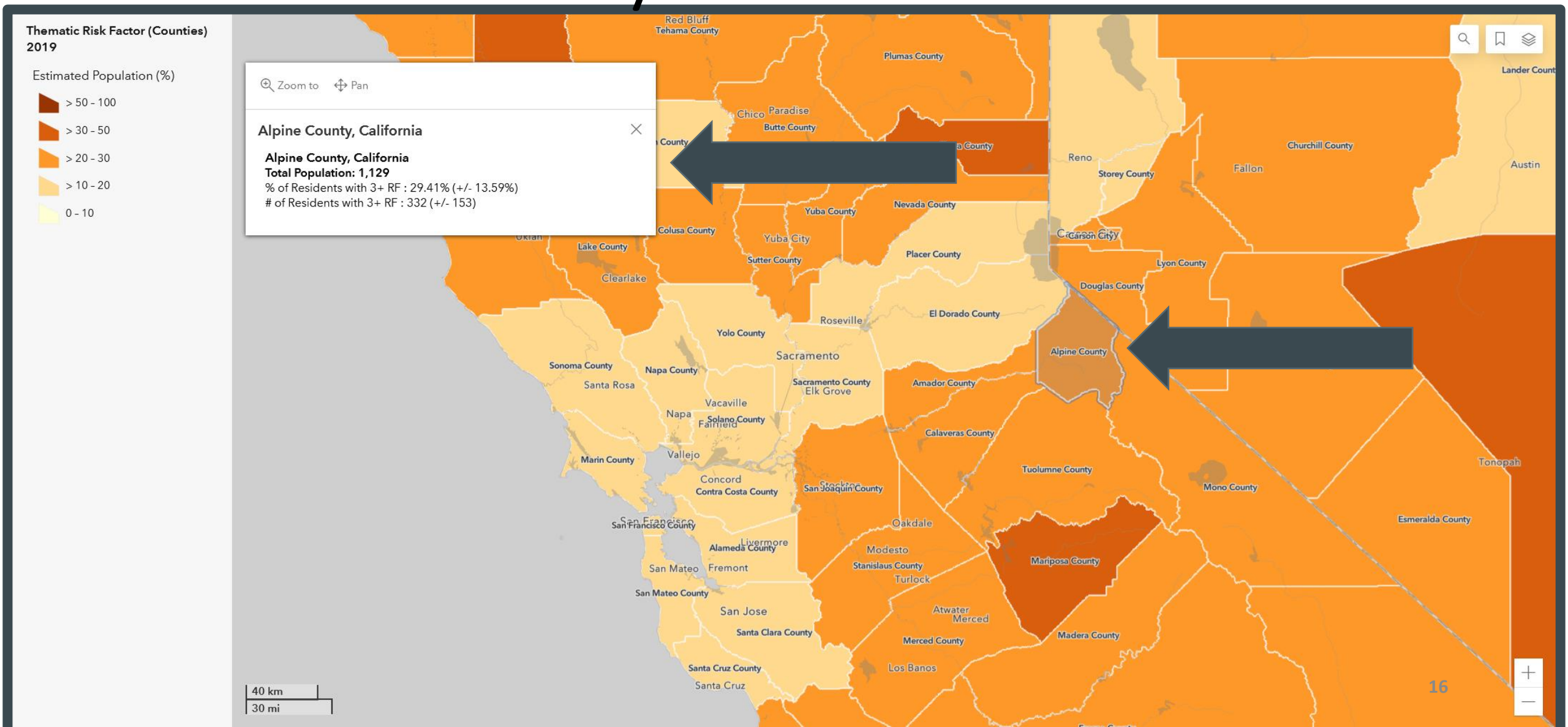


What can these estimates do?

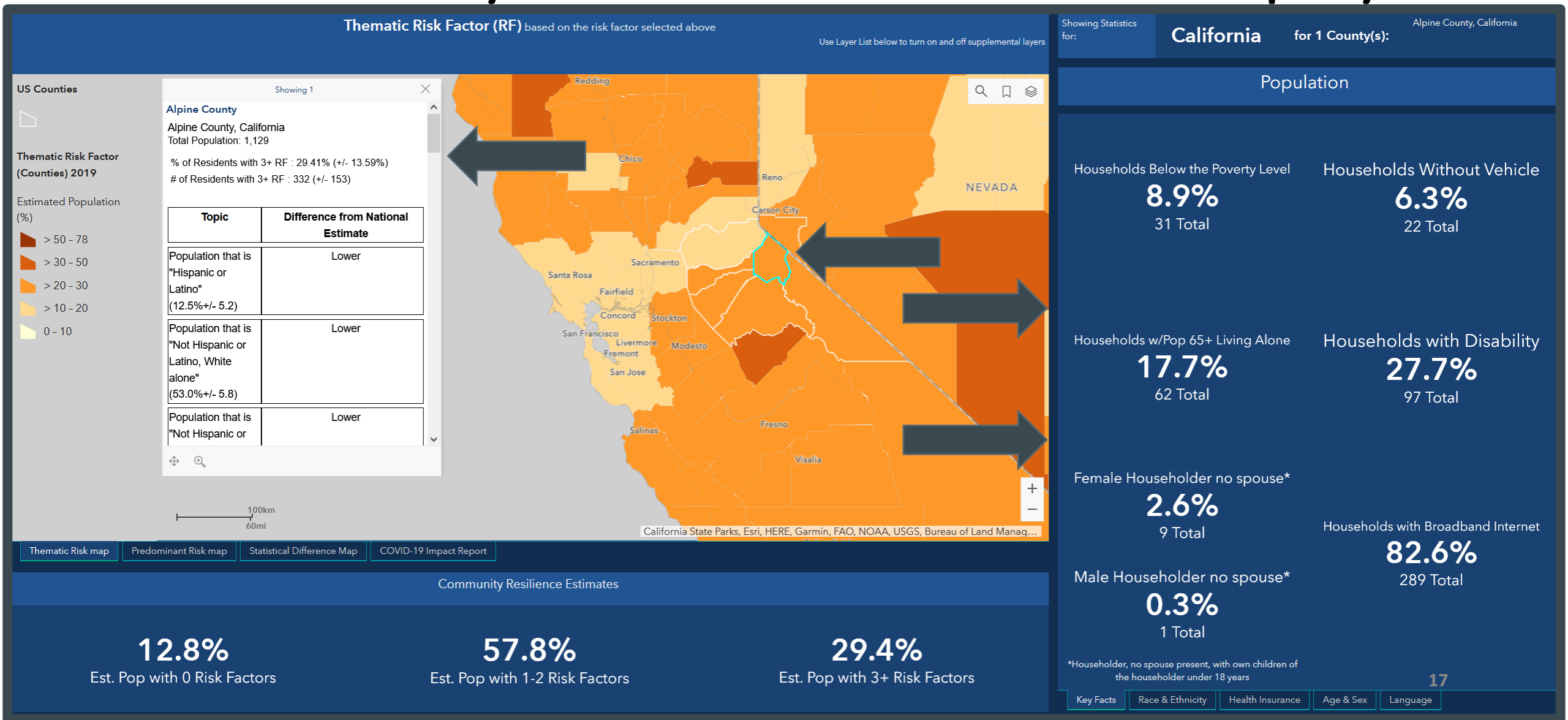
- CRE can answer statistical questions that other indices cannot:
  - What portion of the population has a high-level of socioeconomic vulnerability to disasters (i.e., high-risk population rate) in the United States? What is the rate by region or division?
  - How many people live in vulnerable communities? How about resilient ones?
  - In comparison to small rural, large rural and urban communities, do isolated communities have a higher portion of their population that has a high-level of socioeconomic vulnerability?
  - To what extent is the high-risk population rate of communities in persistent poverty greater than communities not in persistent poverty? How about historically disenfranchised communities?
  - To what extent is the high-risk population rate of toxic communities greater than non-toxic communities?


WA
OR
CA
ID
NV
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
MD
DE
NJ
CT
RI
MA
VT
NH
ME

Community Resilience Estimates



Community Resilience Estimates for Equity



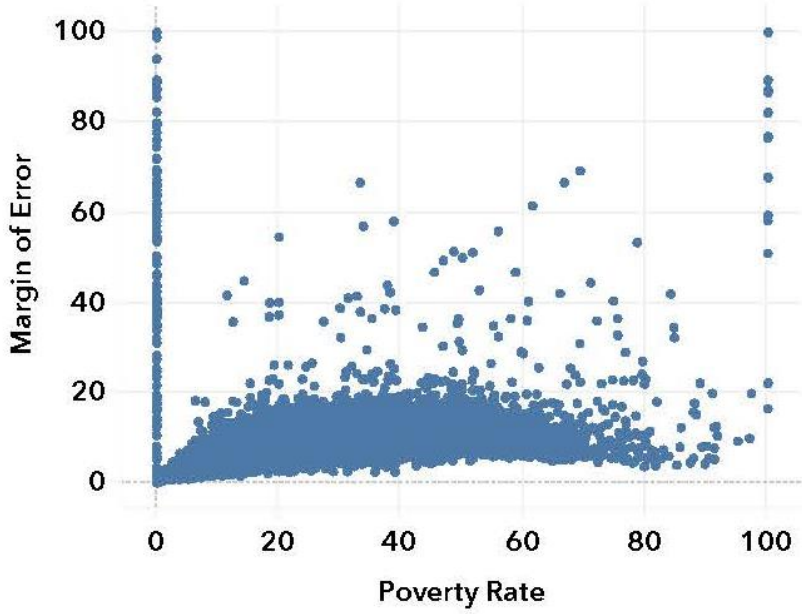
Thank you!

Bethany S. DeSalvo, PhD

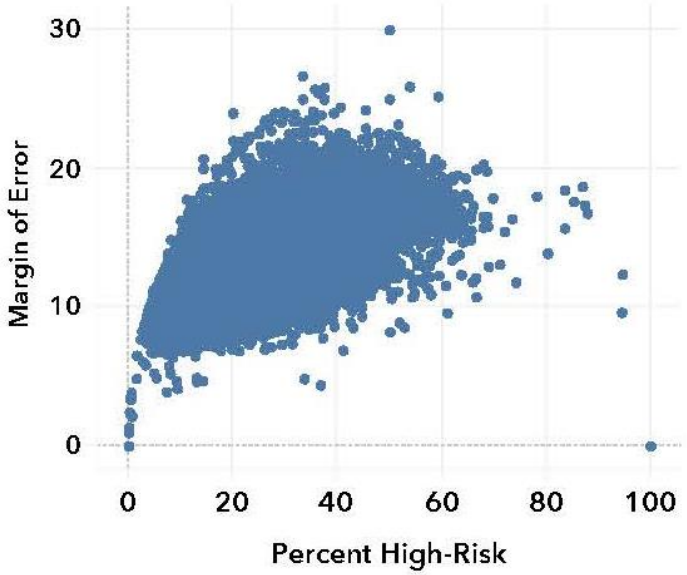
Bethany.desalvo@census.gov

Clustering at extremes

2015-2019 ACS Poverty Rate Estimates and Margins of Error



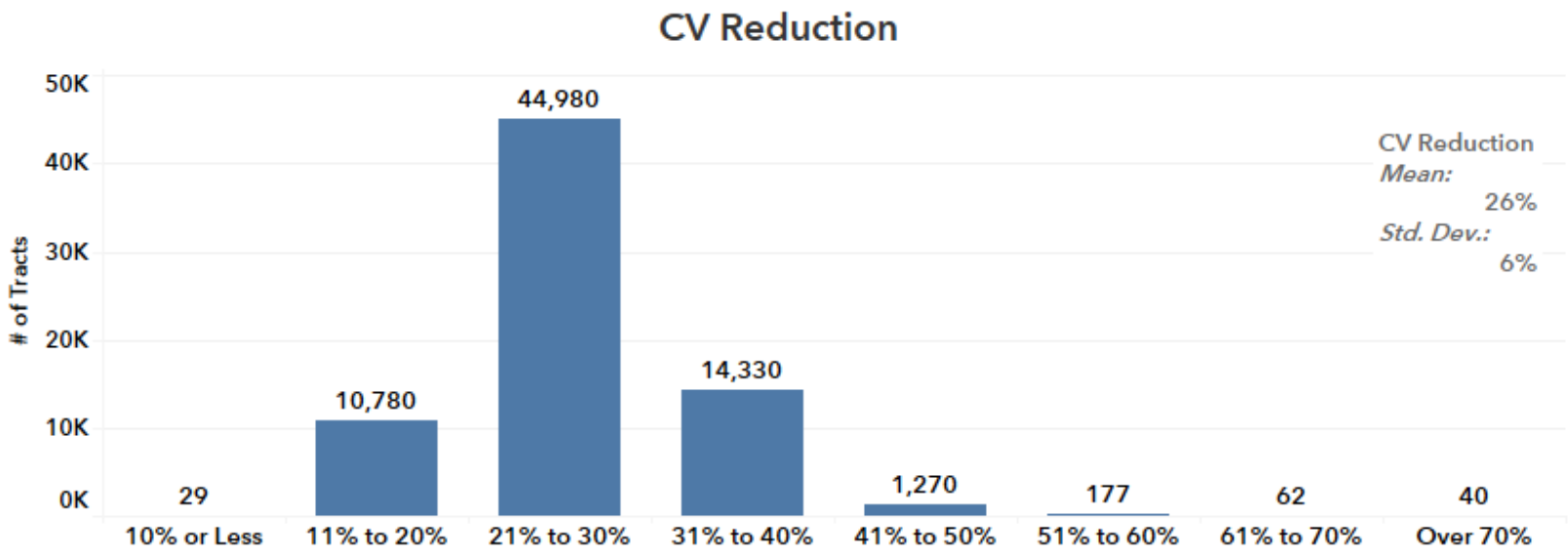
2019 CRE High-Risk Population Rate Estimates and Margins of Error



Source: 2019 CRE and 2015-2019 5-Year American Community Survey

CRE Reduces ACS Error

Description of Percent Reduction in the Relative Error of High-Risk Population for Populated Census Tracts



One-Sample T-Test of the Reduction in Relative Error for High-Risk Population Estimates through Small Area Modeling

TTEST Procedure – Variable: Percent Reduction in Relative Error (N ~ 71,670 Tracts)			
		90% Confidence Level	
Mean	25.95%	25.91%	25.99%
Standard Deviation	6.20%	6.17%	6.23%
	D.F.	T Value	Pr > |t|
	7,242	65,020	2.15

United States Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Source: 2019 CRE Connected to Restricted Complete 2019 1-Year ACS Estimates